AFRL-VA-WP-TP-2002-321

OPTIMAL CONTROL PROBLEMS ON RIEMANNIAN MANIFOLDS: THEORY AND APPLICATIONS

Ram Venkataraman, Raymond Holsapple, and David Doman

AUGUST 2002

20021016 247

AIR VEHICLES DIRECTORATE
AIR FORCE RESEARCH LABORATORY
AIR FORCE MATERIEL COMMAND
WRIGHT-PATTERSON AIR FORCE BASE, OH 45433-7542

REPORT DOCUMENTATION PAGE		Form Approved OMB No. 0704-0188

The public reporting burden for this collection of information is estimated to average 1 hour per response, including the time for reviewing instructions, searching existing data sources, gathering and maintaining the data needed, and completing and reviewing the collection of information. Send comments regarding this burden estimate or any other aspect of this collection of information, including suggestions for reducing this burden, to Department of Defense, Washington Headquarters Services, Directorate for Information Operations and Reports (0704-0188), 1215 Jefferson Davis Highway, Suite 1204, Arlington, VA 22202-4302. Respondents should be aware that notwithstanding any other provision of law, no person shall be subject to any penalty for failing to comply with a collection of information if it does not display a currently valid OMB control number. **PLEASE DO NOT RETURN YOUR FORM TO THE ABOVE ADDRESS.**

1. REPORT DATE *(DD-MM-YY)*	2. REPORT TYPE	3. DATES COVERED *(From - To)*
August 2002	Journal Article Preprint	05/16/02 – 08/16/02

4. TITLE AND SUBTITLE	5a. CONTRACT NUMBER
OPTIMAL CONTROL PROBLEMS ON RIEMANNIAN MANIFOLDS: THEORY AND APPLICATIONS	In-house
	5b. GRANT NUMBER
	5c. PROGRAM ELEMENT NUMBER N/A

6. AUTHOR(S)	5d. PROJECT NUMBER N/A
Ram Venkataraman and Raymond Holsapple (Texas Tech) David Doman (AFRL/VACA)	5e. TASK NUMBER N/A
	5f. WORK UNIT NUMBER N/A

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)		8. PERFORMING ORGANIZATION REPORT NUMBER
Dept. of Mathematics and Statistics Texas Tech University Lubbock, TX 79409	Control Theory Optimization Branch (AFRL/VACA) Control Sciences Division Air Vehicles Directorate Air Force Research Laboratory, Air Force Materiel Command Wright-Patterson AFB, OH 45433-7542	AFRL-VA-WP-TP-2002-321

9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)	10. SPONSORING/MONITORING AGENCY ACRONYM(S)
Air Vehicles Directorate Air Force Research Laboratory Air Force Materiel Command Wright-Patterson Air Force Base, OH 45433-7542	AFRL/VACA
	11. SPONSORING/MONITORING AGENCY REPORT NUMBER(S) AFRL-VA-WP-TP-2002-321

12. DISTRIBUTION/AVAILABILITY STATEMENT
Approved for public release; distribution is unlimited.

13. SUPPLEMENTARY NOTES
Journal article submitted September 26, 2002 for publication in *SIAM Journal of Control*.

14. ABSTRACT
Any air vehicle can be thought of as evolving on Riemannian manifold with the total kinetic energy as the metric. In this paper, we first derive first-order necessary conditions for a Bolza-type optimal control problem on a Riemannian manifold. Then we apply this theory to a rotating rigid body, by obtaining expressions for the Riemannian connection and curvature tensor via Cartan's formalism.

15. SUBJECT TERMS
shooting methods, two-point-boundary-value problem, trajectory generation

16. SECURITY CLASSIFICATION OF:			17. LIMITATION OF ABSTRACT:	18. NUMBER OF PAGES	19a. NAME OF RESPONSIBLE PERSON (Monitor)
a. REPORT Unclassified	b. ABSTRACT Unclassified	c. THIS PAGE Unclassified	SAR	32	David B. Doman 19b. TELEPHONE NUMBER *(Include Area Code)* (937) 255-8451

Standard Form 298 (Rev. 8-98)
Prescribed by ANSI Std. Z39-18

HES&S 31-15093-1

Optimal Control Problems on Riemannian Manifolds: Theory and Applications

Ram Venkataraman,[*] Raymond Holsapple[†]

Department of Mathematics and Statistics

Texas Tech University, Lubbock, TX 79409 USA

David Doman[‡]

U.S. Air Force Research Laboratory

Wright-Patterson Air Force Base, Ohio 45433-7531

August 23, 2002

Abstract

Any air-vehicle can be thought of as evolving on Riemannian manifold with the total kinetic energy as the metric. In this paper, we first derive first-order necessary conditions for a Bolza-type optimal control problem on a Riemannian manifold. Then we apply this theory to a rotating rigid body, by obtaining expressions for the Riemannian connection and curvature tensor via Cartan's formalism.

The upshot of this work is the derivation of first-order necessary conditions for the trajectory-planning problem that are singularity-free, and considerably simpler that those obtained by using an Euler angles representation.

We next apply a new numerical solution technique called the "Modified Simple Shooting Method", to the resulting two-point boundary value problem. We demonstrate that by using the new necessary conditions and the numerical method, one can solve the trajectory planning problem extremely fast and on-line implementations become feasible.

ASC -02- 2325

1 Introduction

Any air-vehicle can be thought of as evolving on Riemannian manifold with the total kinetic energy as the Riemannian metric. If we consider the aerodynamic forces and moments to be inputs, then the state variables are (Q, ω, b, v) where Q is the orientation of the vehicle with respect to a fixed co-ordinate axis; ω is the body angular velocity; b is the position of the center-of-mass from the origin of the fixed co-ordinate axis; and v is the velocity of the vehicle in the body axes co-ordinates. The state space is therefore, $TSE(3)$ the tangent bundle of the group $SE(3)$. The on-line trajectory planning problem (in case of actuator failures) for aircraft is thus naturally posed on a Riemannian manifold rather than a Euclidean space. In this paper, we first study optimal control problems on parallelizable Riemannian manifolds. Then we consider the solution of the resulting two-point boundary value problems with a new numerical technique called the modified simple shooting method.

In section 2, we derive first-order necessary conditions for the extremals of a Bolza-type optimal control problem on parallelizable Riemannian manifolds using calculus of variations. Previous work in this area was done by P. Crouch, M. Camarinha and F. Silva Leite in a series of papers [1, 2, 3]. Their work and ours differ in the nature of variations considered, and therefore the necessary conditions obtained in this paper are different from those of P. Crouch, M. Camarinha and F. Silva Leite.

In [4], Sussmann tackled the problem of generalizing the minimum principle (or equivalently the maximum principle by considering the negative of the cost function), to manifolds (without any affine-connection structure), by developing the co-ordinate free maximum principle. However, when this principle is applied to an air-vehicle problem, one employs local co-ordinates and the equations reduce to the necessary conditions for an optimal control problem in co-ordinates. In contrast, we make use of the Riemannian connection and employ frame co-ordinates (we assume the manifold to be parallelizable), that yield global equations. Thus the method is frame dependent rather than invariant. This idea can also be seen in the work of P. Crouch, M. Camarinha and F. Silva Leite [1, 2, 3]. In subsection 3.1, we show that one can obtain the Noakes, Heinzinger and Paden formula for cubic splines on Riemannian manifolds [5] by specializing Theorem 2.3.

Finally, we apply the theory to the path-planning problem for a rotating rigid body as an optimal control problem on TSO(3). We obtain the first order necessary conditions for a rigid body without the use of local co-ordinates, but rather frame co-ordinates. In Subsection 3.2, we

2

present a numerical experiment where these equations are solved using a modified simple shooting method.

2 Optimal Control on Riemannian Manifolds

In this section, we derive necessary conditions for Bolza-type optimal control problems on Riemannian manifolds. Previous work in this area was done by P. Crouch, M. Camarinha and F. Silva Leite in a series of papers [1, 2, 3]. However, our work differs in the nature of variations considered, and in this respect more closely resembles the work of L. Noakes, G. Heinzinger and B. Paden [5]. More precisely, suppose that M is a Riemannian manifold with ∇ as the Levi-Civita connection. Suppose further that $(q, V)(t)$, $t \in [t_0, t_f]$ denotes the optimal solution as a curve in TM. Here, $q(t) \in M$, and $V(t) \in T_{q(t)}M$, $t \in [t_0, t_f]$. Now consider a one-parameter family of variations given by $(q, V)(t, \sigma)$, $t \in [t_0, t_f]$, $\sigma \in (-\epsilon, \epsilon)$, $\epsilon > 0$. For optimal control problems with both initial and final states (that is, $(q, V)(t_0)$ and $(q, V)(t_f)$) are prescribed, the admissible set of variations for Crouch et al. are those for which the ordinary derivatives $\frac{\partial q}{\partial \sigma}$ and $\frac{\partial V}{\partial \sigma}$ vanish at t_0 and t_f. In this paper however, we consider $\frac{\partial q}{\partial \sigma}$, and the total or covariant derivative of the velocity $\nabla_{\frac{\partial q}{\partial \sigma}} V$ to vanish at t_0 and t_f. The justification for this, is a lemma proved by Noakes, Heinzinger and Paden [5]. Therefore, the necessary conditions obtained in this work are different from those of P. Crouch, M. Camarinha and F. Silva Leite.

2.1 Parallelizable manifolds and Cartan Formalism

This subsection details the mathematical background in differential geometry used in the rest of the paper. There are several good sources for the material such as Frankel [6] and Petersen [7]. First, let us consider the concept of *parallelizable* manifolds.

Definition 2.1 *(Parallelizable Manifolds) [8] If there exists a set of n smooth vector fields on a manifold M of dimension n that are linearly independent at each point, then the manifold M is said to be parallelizable.*

The above set of n linearly independent vector fields is referred to as a *field of co-ordinate frames* or briefly as a *frame*. It is well known that the condition of being parallelizable is very special for a manifold. For example, among the spheres S^n, $n = 1, 2, \cdots$, only S^1, S^3 and S^7 are parallelizable [8]. However, the following theorem states that all Lie groups are parallelizable.

3

Theorem 2.1 *[8] Let G be a Lie group and $\mathcal{G}$ the tangent space at the identity e. Then each $X_e \in \mathcal{G}$ determines uniquely a C^∞ vector field X on G which is invariant under left translations. In particular, G is parallelizable.*

In the case of air vehicles, the configuration space as mentioned earlier is the Lie Group $SE(3)$. However the manifold on which the air vehicle dynamics can be described is $TSE(3)$ the tangent bundle of $SE(3)$. So we need the following lemma:

Lemma 2.1 *Let M be a parallelizable manifold. Then TM is parallelizable.*

Proof

This follows from the observation that locally TM is a product manifold $U \times \mathbb{R}^n$ where U is a co-ordinate neighborhood of M. Now $\mathbb{R}^n$ is parallelizable as it is a Lie Group. Let $\{X_1, \cdots, X_n\}$ be a frame on M. Then we can define vector fields $F_{ij} = (X_i, E_j)$, $i, j = 1, \cdots, n$ where E_j are the co-ordinate vector fields for $\mathbb{R}^n$, that form a frame on TM. ∎

Now, let M be a manifold with a symmetric, positive definite, and bilinear form defined on $T_q M$ where $q \in M$, denoted by $< \cdot, \cdot >_q$. Such a form is usually referred to as *Riemannian metric* and we assume it to be smooth for each $q \in M$ and C^1 as a function of q. A Riemannian metric defines an linear isomorphism $\Sigma : T_q M \to T_q^* M$ where $q \in M$, by:

$$(\Sigma(v))(w) := < v, w >_q; \quad v, w \in T_q M. \tag{1}$$

As Σ is an isomorphism, we also have the inverse map $\Sigma^{-1} : T_q^* M \to T_q M$ where $q \in M$.

Let $\Psi(M)$ denote the set of all C^∞ vector fields on M. If $X \in \Psi(M)$, then the *directional derivative* of a C^∞ function f on M in the direction of X is another C^∞ function denoted by $X(f(q))$. If $X, Y \in \Psi(M)$, then one can define another vector field called the *Lie bracket* of X and Y by:

$$[X, Y](f) := X(Y(f)) - Y(X(f)). \tag{2}$$

If $X, Y, Z \in \Psi(M)$, then an affine *connection* is an operator $\nabla : \Psi(M) \times \Psi(M) \to \Psi(M)$, satisfying:(a) $\nabla_X(aY + bZ) = a\nabla_X Y + b\nabla_X Z$; $a, b \in \mathbb{R}$, (b) $\nabla_{(aX+bY)}Z = a\nabla_X Z + b\nabla_Y Z$, and (c) $\nabla_X(fY) = X(f)Y + f\nabla_X Y$; $f \in C^\infty(M)$.

It is well known that on a Riemannian manifold M, there exists a unique, affine connection $\nabla : \Psi(M) \times \Psi(M) \to \Psi(M)$ on M that is (a) symmetric (also called *torsion-free*), i.e. $\nabla_X Y - \nabla_Y X = [X, Y]$, $X, Y \in \Psi(M)$ and (b) compatible with the Riemannian metric, that is, $X < Y, Z >$

$= \ <\nabla_X Y, Z> + <Y, \nabla_X Z>, \ X, Y, Z \in \Psi(M)$ [9]. Such a connection is called a *Levi-Civita* connection on M.

Assume M to be parallelizable and consider a frame $\{E_i\}$ and a dual co-frame $\{\theta^i\}$, $i = 1, 2, \cdots n$ with $\theta^i(E_j) = \delta^i_j$. Then the Levi-Civita connection is defined by the functions $\omega^k_{ij}(q)$ in the equation:

$$\nabla_{E_i} E_j = \omega^k_{ij} E_k, \quad 1 \leq i, j \leq n. \tag{3}$$

If the vector field $X \in \Psi(M)$ is defined by $X = x^i E_i$, then we have:

$$\begin{aligned}
\nabla_X E_j &= x^i \nabla_{E_i} E_j \\
&= x^i \omega^k_{ij} E_k \\
&= \left(\omega^k_{ij} E_k \otimes \theta^i \right) (x^i E_i),
\end{aligned}$$

where $\otimes$ is the mixed tensor product of a vector and co-vector. Thus if $\eta, \xi \in T^*_q M$ and $v, w \in T_q M$, then $v \otimes \eta$ acts on (ξ, w) by:

$$(v \otimes \eta)(\xi, w) = \xi(v)\eta(w).$$

Similarly, one can define a tensor product of two forms

$$(\eta \otimes \xi)(v, w) = \eta(v)\xi(w).$$

This tensor product is not a two-form. The wedge product of two one-forms $\eta, \xi \in T^*M$ is a two-form defined by:

$$\eta \wedge \xi = \eta \otimes \xi - \xi \otimes \eta.$$

If we denote $\omega^k_j := \omega^k_{ij} \theta^i$, then we can define the $n \times n$ matrix ω of connection one-forms by:

$$\omega := (\omega^k_j). \tag{4}$$

We can then write compactly:

$$\nabla E_j = E_k \otimes \omega^k_j. \tag{5}$$

Thus by computing the connection matrix ω, one can obtain the Levi-Civita connection on M. For a *co-ordinate frame*, that is, a frame that results from a local co-ordinate chart, the connection coefficients are symmetric in the lower indices: $\omega^k_{ij} = \omega^k_{ji}$. This is due to the fact that a co-ordinate frame has the property that the Lie-bracket of the co-ordinate vector fields, $[E_i, E_j] = 0$, $1 \leq i, j \leq$

$$\begin{array}{ccc}
\text{TTM} & \xrightarrow{\;K\;} & \text{TM} \\
\Big\downarrow \scriptstyle\pi_{TM} & & \Big\downarrow \scriptstyle\pi_M \\
\text{TM} & \xrightarrow{\;\pi_M\;} & \text{M}
\end{array}
\qquad
\begin{array}{ccc}
((q,v),(w,r)) & \xrightarrow{\;K\;} & (q, r + (\omega(v))(w)) \\
\Big\downarrow \scriptstyle\pi_{TM} & & \Big\downarrow \scriptstyle\pi_M \\
(q,v) & \xrightarrow{\;\pi_M\;} & q
\end{array}$$

Figure 1: Use of the connection to map vectors on TTM to TM.

n. However, we shall not use co-ordinate frames in our study of the rigid body in Section 3. If $Y \in \Psi(M)$ with $Y = y^j E_j$ in frame co-ordinates then:

$$\begin{aligned}
\nabla_X Y &= \nabla_{x^i E_i} y^j E_j \\
&= x^i E_i(y^j) E_j + x^i y^j \omega_{ij}^k E_k \\
&= x^i E_i(y^j) E_j + \omega_j^k(X) y^j E_k \\
&= x^i E_i(y^j) E_j + (\omega(X))(Y)
\end{aligned}$$

One can use a connection structure to map a vector field on TTM to one on TM. To be precise, let $\pi_M : TM \to M$; $\pi(q, V) = q$ be the trivial projection map. This map is a smooth map in the topologies of TM and M. Similarly, we have the trivial projection, $\pi_{TM} : TTM \to TM$; $\pi((q, v), (w, r)) = (q, v)$. On the other hand, we have differential of projection map on M given by, $d\pi_M : TTM \to TM$; $d\pi((q, v), (w, r)) = (q, w)$. The kernel of $d\pi$ at every $q \in M$ defines the so-called *vertical subspace* of $T_{(q,v)}TM$. In co-ordinates, this subspace consists of elements of the type $((q, v), (0, r))$. A horizontal subspace of $T_{(q,v)}TM$ is obtained using the connection ∇. Consider the vector bundle morphism in Figure 1. The kernel of the map K is a *horizontal subspace* of $T_{(q,v)}TM$. In co-ordinates, this subspace consists of elements of the type $((q, v), (w, -\omega(v, w)))$. These two subspaces lead to a splitting of $T_{(q,v)}TM$ [10]. We do not need this structure in this paper, except to note that this construction yields us a way of mapping a vector field on TTM to one on TM via the connection map K. We use this mapping to define a second-order control system on M in the next subsection.

The following propositions due to E. Cartan formalizes the computations:

Proposition 2.1 *[7] Define $d\theta = (d\theta^1, \cdots, d\theta^n)$. Then there is a unique matrix of 1-forms $\omega = \left(\omega^i_{\,j} \right)$ such that:*

$$d\theta = -\omega \wedge \theta \tag{6}$$

$$d\theta^i = -\omega^i_{\,j} \wedge \theta^j \tag{7}$$

6

and

$$\omega^i_j = -\omega^j_i. \tag{8}$$

These are known as the first structural equations. The first equation encodes the fact that the connection is torsion-free and the second equation is due to the compatibility of the connection with the metric. In section 3, we use the above proposition to compute the Levi-Civita connection for a rotating rigid body. The left hand side of the equation (7) can be calculated using the following theorem:

Theorem 2.2 *[6] Let α be a one-form on a smooth manifold M. Let $q \in M$ and $X_q, Y_q \in T_q M$. Extend these vectors in any smooth way to be vector fields near q. Then*

$$d\alpha(X_q, Y_q) = X_q(\alpha(Y_q)) - Y_q(\alpha(X_q)) - \alpha([X,Y]_q) \tag{9}$$

Let X and Y be two vector fields on M. Then the vector fields

$$R(X,Y)E_j := \nabla_X \nabla_Y E_j - \nabla_Y \nabla_X E_j - \nabla_{[X,Y]} E_j, \quad j = 1, \cdots, n \tag{10}$$

define the *curvature transformation* on M for the pair X, Y (please note that our definition follows those of Petersen [7], Helgason [11], Boothby [8], Klingenberg [10] and Frankel [6] but is the negative of Do Carmo's definition [9]). The linear transformation $R(X,Y) : \Psi(M) \to \Psi(M)$ has the matrix

$$R(X,Y)^i_j = R^i_{jkl} X^k Y^l.$$

Consequently, R^i_{jkl} define a mixed $(1,3)$ tensor called the *curvature tensor*. It is well known that the curvature tensor satisfies the properties listed in the following proposition.

Proposition 2.2 *[7] Let X, Y, Z, W be vector fields on M. Then:*

1. $<R(X,Y)Z, W> = - <R(Y,X)Z, W> = - <R(X,Y)W, Z>$;

2. $<R(X,Y)Z, W> = <R(Z,W)X, Y>$;

3. (Bianchi's first identity) $R(X,Y)Z + R(Z,X)Y + R(Y,Z)X = 0$; and

7

4. *(Bianchi's second identity)* $(\nabla R)(Z, X, Y, W) + (\nabla R)(X, Y, Z, W) + (\nabla R)(Y, Z, X, W) = 0$
   where

$$(\nabla R)(Z, X, Y, W) := (\nabla_Z R)(X, Y)W = \nabla_Z(R(X, Y)W) - R(\nabla_Z X, Y)W$$
$$-R(X, \nabla_Z Y)W - R(X, Y)\nabla_Z W$$

is the covariant differentiation of the curvature tensor.

The first two identities in Proposition 2.2 are used in the next section. Cartan showed that the curvature tensor can be obtained from the connection matrix:

Proposition 2.3 *[7] The equations*

$$\Omega = d\omega + \omega \wedge \omega \tag{11}$$

$$\Omega^i_{\ j} = d\omega^i_{\ j} + \omega^i_{\ k} \wedge \omega^k_{\ j} \tag{12}$$

define a skew-symmetric matrix of 2-forms that is related to the curvature tensor via

$$R(X, Y)E_j = \Omega^i_{\ j}(X, Y) \otimes E_i \tag{13}$$

The above proposition is used in Section 3 to compute the curvature tensor for the rigid body rotation problem.

2.2 Derivation of first-order necessary conditions

Let M be a parallelizable Riemannian manifold as in the last sub-section. If $c : [0, 1] \to M$ is a differentiable curve on M, and $X : M \to TM$ is a differentiable vector field, then the co-variant derivative of X along $c(\cdot)$ is defined in the standard manner (see Boothby [8] or Do Carmo [9]) to be $\frac{DX}{dt} = \nabla_{\dot{c}(t)} X(t)$, $t \in (0, 1)$. Let $\{E_1, \cdots, E_n\}$ be a frame of vector fields and let $\{\theta^1, \cdots, \theta^n\}$ be a frame of co-vector fields on M, so that $\theta^i(E_j) = \delta^i_j$; $1 \le i, j \le n$.

We define a *control system* on TM by a second-order vector field $F : TM \times \mathbb{R}^m \to TTM$ defined as follows. If $\pi : TM \to M$ denotes the projection operator, then a second-order vector field is one that satisfies $d\pi \circ F_{(q,V)} = (q, V)$, where $q \in M$, and $V \in T_q M$. If $u \in \mathbb{R}^m$, then F is locally defined by $((q, V), u) \mapsto ((q, V), (V, \tilde{f}(q, V, u))$ or in other words, by the differential equations:

$$\dot{q} = V,$$
$$\dot{V} = \tilde{f}(q, V, u)$$

8

at the point $(q, V) \in TM$. Using the Levi-Civita connection on M (as outlined in the previous subsection), we can write the above system as one on TM described by the equations:

$$
\begin{aligned}
\dot{q} &= V = V^i E_i, \\
\frac{DV}{dt} &= f(q, V, u) = f^i(q, V, u)E_i.
\end{aligned}
\tag{14}
$$

The conditions on the vector field $\tilde{f}$ and by extension the function f will be set forth in our theorem on necessary conditions. Such a set of equations is useful in describing the equations of motion of an air vehicle that is subject to aerodynamic forces and moments (as well as gravity) that depend on its orientation with respect to its velocity vector, its height above sea level, current speed and the deflections of its control surfaces.

Now, let $q_0, q_1 \in M$, $V_0 \in T_{q_0}M$ and $V_1 \in T_{q_1}M$. Consider the space $C_2[0, 1]$ of twice-differentiable maps $q : [t_0, t_f] \to M$ that satisfy Equations (14), where $t_f > t_0$, $q(t_0) = q_0$, $q(t_f) = q_f$, $\dot{q}(t_0) = V_0$ and $\dot{q}(t_f) = V_f$. Then along one such map $q(\cdot)$ the control system takes the form:

$$
\begin{aligned}
\dot{q}(t) &= V(t), \tag{15} \\
\frac{DV}{dt} &= f(q(t), V(t), u(t)), \tag{16}
\end{aligned}
$$

where $u(\cdot) \in C^m[t_0, t_f]$, the (m-vector valued) space of continuous functions. Suppose that one is required to find a function $u(\cdot)$ such that the above boundary conditions are satisfied by $q(\cdot)$ while minimizing:

$$
J(u(\cdot)) = \int_{t_0}^{t_f} L(q(t), V(t), u(t))dt.
\tag{17}
$$

We need the following standard construction to describe the notion of variations of a curve (we use the same notation here as in Crouch, Camarinha and Silva Leite [2]). Let $(t, \sigma) \to q(t, \sigma)$, $t \in [t_0, t_f]$ and $\sigma \in (-\epsilon, \epsilon)$, $\epsilon > 0$, be a parametrized family of curves satisfying

$$
\begin{aligned}
q(t, 0) &= q(t) \\
q(t_0, \sigma) &= q_0 \\
q(t_f, \sigma) &= q_f \\
\dot{q}(t_0, \sigma) &= V_0 \\
\dot{q}(t_f, \sigma) &= V_f
\end{aligned}
\tag{18}
$$

For $V(t) \in T_{q(t)}M$ and $p_1(t), p_2(t) \in T_{q(t)}^*M$, we define the associated variations:

$$
V(t, \sigma) = V^i(t, \sigma)E_i(q(t, \sigma)) \in T_{q(t, \sigma)}M
$$

$$p_1(t,\sigma) = p_{1_i}(t,\sigma)\theta^i(q(t,\sigma)) \in T^*_{q(t,\sigma)}M$$

$$p_2(t,\sigma) = p_{2_i}(t,\sigma)\theta^i(q(t,\sigma)) \in T^*_{q(t,\sigma)}M \tag{19}$$

Define the variational vector fields

$$W(t) = \delta q(t) := \tfrac{\partial q}{\partial \sigma}(t,0) \in T_{q(t)}M,\ t \in [0,T], \tag{20}$$

$$W(0) = W(T) = 0; \qquad\qquad \text{and} \tag{21}$$

$$\delta V(t) := \tfrac{DV}{d\sigma}(t,0) \in T_{q(t)}M,\ t \in [0,T]. \tag{22}$$

The variations in the input are denoted by $u(t,\sigma) \in \mathbb{R}^m$ with

$$\delta u(t) := \frac{\partial u}{\partial \sigma}(t,0) \in \mathbb{R}^m. \tag{23}$$

In the following, any quantity that is described with the second variable σ suppressed, should be construed as having $\sigma = 0$ (so, $q(t) = q(t,0)$).

Then we have the following lemma proved by Noakes, Heinzinger and Paden.

Lemma 2.2 *[5]* $\frac{DV}{d\sigma}(t_0,\sigma) = 0$ *and* $\frac{DV}{d\sigma}(t_f,\sigma) = 0$ *for all* $\sigma \in (-\epsilon, \epsilon)$.

Crouch, Camarinha and Silva Leite set the ordinary derivatives $\frac{\partial V}{\partial \sigma}(t_0,\sigma)$, $\frac{\partial V}{\partial \sigma}(t_f,\sigma)$ equal to 0 and thus obtain different results from our work. We need the following simple lemmas in the proof of the main theorem (Theorem 2.3).

Lemma 2.3

$$\int_{t_0}^{t_f} p_1(t)(\frac{D\dot{q}}{d\sigma}(t))dt = -\int_{t_0}^{t_f} \frac{Dp_1}{dt}(\delta q(t))dt \tag{24}$$

Proof We have $\frac{D\dot{q}}{d\sigma}(t) = \frac{DW}{dt}(t)$ because the Levi-Civita connection is symmetric and $[\frac{\partial}{\partial t}, \frac{\partial}{\partial \sigma}] = 0$ (for full details on the construction that leads to this conclusion, please refer to Noakes, Heinzinger and Paden [5]). Since the Levi-Civita connection is compatible with the Riemannian metric, we have:

$$
\begin{aligned}
\int_{t_0}^{t_f} \frac{d}{dt}p_1(\delta q)\,dt &= \int_{t_0}^{t_f} \frac{d}{dt} <\Sigma^{-1}p_1, \delta q>dt \\
&= \int_{t_0}^{t_f} (<\frac{D}{dt}\Sigma^{-1}p_1, \delta q> + <\Sigma^{-1}p_1, \frac{D}{dt}\delta q>)\,dt \\
&= \int_{t_0}^{t_f} (\frac{Dp_1}{dt}(\delta q) + p_1(\frac{D}{dt}\delta q))\,dt,
\end{aligned}
$$

10

that leads to the "integration by parts" formula:

$$\int_{t_0}^{t_f} p_1(\frac{D}{dt}\delta q)\, dt = p_1(\delta q)|_{t_0}^{t_f} - \int_{t_0}^{t_f} \frac{Dp_1}{dt}(\delta q)\, dt.$$

Therefore:

$$\begin{aligned}
\int_{t_0}^{t_f} p_1(t)(\frac{D\dot{q}}{d\sigma}(t))dt &= \int_{t_0}^{t_f} p_1(t)(\frac{DW}{dt}(t)) \\
&= p_1(W)|_{t_0}^{t_f} - \int_{t_0}^{t_f} \frac{Dp_1}{dt}(\delta q)dt,
\end{aligned}$$

on integration by parts. Now the result follows by noting that $W(t_0) = W(t_f) = 0$. ∎

Lemma 2.4

$$\int_{t_0}^{t_f} p_2(t)\frac{D}{d\sigma}\frac{DV}{dt}dt = \int_{t_0}^{t_f} \left((\Sigma R(\Sigma^{-1}p_2, V)V)\delta q - \frac{Dp_2}{dt}(\delta V(t)) \right) dt \qquad (25)$$

Proof First we note that:

$$\int_{t_0}^{t_f} p_2(t)\frac{D}{d\sigma}\frac{DV}{dt}dt = \int_{t_0}^{t_f} p_2(t)(R(W,V)V + \frac{D}{dt}\frac{DV}{d\sigma})dt,$$

by the definition of the curvature tensor in Equation (10), and the fact that $[\frac{\partial}{\partial t}, \frac{\partial}{\partial \sigma}] = 0$. By the definition of the linear isomorphism Σ given in Equation (1) and Proposition 2.2, we have:

$$\begin{aligned}
p_2\,(R(W,V)V) &= <\Sigma^{-1}p_2, R(W,V)V> \\
&= <R(V,\Sigma^{-1}p_2)W, V> \\
&= <R(\Sigma^{-1}p_2, V)V, W> \\
&= (\Sigma R(\Sigma^{-1}p_2, V)V)(W)
\end{aligned}$$

Therefore,

$$\int_{t_0}^{t_f} p_2(t)\frac{D}{d\sigma}\frac{DV}{dt}dt = \int_{t_0}^{t_f} \left((\Sigma R(\Sigma^{-1}p_2, V)V)(\delta q) - \frac{Dp_2}{dt}\delta V \right) dt + p_2(\delta V)|_{t_0}^{t_f},$$

on integrating the second term by parts. Now Lemma 2.2 proves the claim. ∎

We need some more notation in order to express the necessary conditions in a compact form. As $\{E_1, \cdots, E_n\}$ forms a global frame of vector fields on M, the Jacobi-Lie brackets of the co-ordinate vector fields $[E_i, E_j]$ can be expressed as a linear combination of $\{E_1, \cdots, E_n\}$:

$$[E_k, E_j] := C_{kj}^i E_i. \qquad (26)$$

11

The coefficients C_{kj}^i are called structure constants.

For $p \in T^*M$, denote:

$$[\omega(f)]^*p := p_i \omega^i_{\ j}\,(f)\theta^j = p_i \omega^i_{kj}\,f^k\theta^j. \tag{27}$$

The motivation for this definition is the fact that, if $F : M \to N$ is a Frechét differentiable mapping between manifolds M and N, then its differential DF maps TM to TN while the pullback $(DF)^*$ maps T^*N to T^*M. They are related as follows: For $q \in M$, suppose $\bar\eta = \bar\eta_i \bar\theta^i \in T^*_{F(q)}N$, $\eta = \eta_i\theta^i \in T^*_q M$ are one-forms related by $\eta = (D_q F)^*(\bar\eta)$ (where, $\{\bar\theta^1, \cdots, \bar\theta^n\}$ and $\{\theta^1, \cdots, \theta^m\}$ form a basis of one-forms in local co-ordinates for $T^*_q M$ and $T^*_{F(q)}N$ respectively). If $v = v^j E_j \in T_q M$, then $\eta(v) = \eta_j v^j = \bar\eta_i (D_q F)^i_j v^j$ for all $v \in T_q M$ so that $\eta = (D_q F)^*(\bar\eta) = \bar\eta_i (D_q F)^i_j \theta^j$.

Similarly denote:

$$\begin{aligned}
[C(f)]^*p &:= p_i C^i_{\ j}\,(f)\theta^j \\
&= p_i C^i_{kj}\,f^k\theta^j. \tag{28}
\end{aligned}$$

The following theorem is the main result of this paper. It establishes the first-order necessary conditions for the curve $(q_0, V_0, u_0)(t)$, $t \in [t_0, t_f]$ to be optimal.

Theorem 2.3 *Suppose that $(q_0, V_0, u_0)(t)$, $t \in [t_0, t_f]$ minimizes the cost function (17), while satisfying Equations (15, 16) and boundary conditions $q(t_0) = q_0$, $q(t_f) = q_f$, $\dot q(t_0) = V_0$ and $\dot q(t_f) = V_f$. Further suppose that*

- *f and L are differentiable functions of their arguments,*

- *the linearized system is controllable at the origin.*

Then there exists one-forms $p_1(t), p_2(t)$, $t \in [0, T]$ and differentiable a.e. on $[0, T]$ such that:

- $$\begin{aligned}
\frac{Dp_1}{dt} &= L_q(q_0, V_0, u_0) + \Sigma R(\Sigma^{-1}p_2, V_0)V_0 - (f^*_q(q_0, V_0, u_0) + \omega^*(f) - C^*(f))p_2 \\
\frac{Dp_2}{dt} &= -p_1 + L_V(q_0, V_0, u_0) - f^*_V(q_0, V_0, u_0)p_2
\end{aligned}$$
  where f denotes the vector field $f(q_0, V_0, u_0)$,

- *$f^*_u(q_0, V_0, u_0)p_2 = L_u(q_0, V_0, u_0)$*

- *the Hamiltonian function*

$$H(q_0, V_0, u_0, p_1, p_2)(t) = L(q_0, V_0, u_0)(t) - p_1(V_0)(t) - p_2(f(q_0, V_0, u_0))(t)$$

is a constant for $t \in [t_0, t_f]$.

The proof of the above theorem will be given shortly. First we make some important remarks. The theorem that yields the existence of the one-forms $p_1(\cdot), p_2(\cdot)$ is the *Lagrange Multiplier* theorem:

Theorem 2.4 *[12] Let $f : X \to \mathbb{R}$ and $H : X \to Z$ be continuously Fréchet differentiable, where X, Z are Banach spaces. Suppose that f has a local extremum under the constraint $H(x) = 0$ at the regular point x_0. Then, there exists an element $z_0^* \in Z^*$ (Z^* is the dual space of Z), such that the functional $f(x) + z_0^* H(x)$ is stationary at x_0, i.e., $f'(x_0) + z_0^* H'(x_0) = 0$.*

The term *regular point* in the above theorem is defined below:

Definition 2.2 *Let T be a continuously Fréchet differentiable map from an open set D in a Banach space X into a Banach space Y. If $x_0 \in D$ is such that $T'(x_0)$ maps X onto Y, the point x_0 is said to be a regular point of the transformation T.*

Typically, theorems showing first order necessary conditions are proved as follows [12]. Let the system be given by:

$$\dot{x}(t) = f(x, u), \quad x(t_0), x(t_f) \text{ fixed,}$$

and the functional to be minimized be

$$J(x(\cdot), u(\cdot)) = \int_{t_0}^{t_f} l(x, u) dt,$$

where $x \in \mathbb{R}^n$ and $u \in \mathbb{R}^m$; that is, we are working in some local co-ordinates. The differential equation is equivalent to the integral equation:

$$x(t) - x(t_0) - \int_{t_0}^{t} f(x(\tau), u(\tau)) d\tau = 0;$$

that is,

$$A(x, u) = 0,$$

where $A : X \times U \to X$. If $X = C^n[t_0, t_f]$ and $U = C^m[t_0, t_f]$, then, the Fréchet differential of A exists; is continuous under our assumptions; and is given by:

$$\delta A(x, u; h, v) = h(t) - \int_{t_0}^{t} f_x h(\tau) d\tau - \int_{t_0}^{t} f_u v(\tau) d\tau$$

for $(h, v) \in X \times U$. Luenberger [12] shows that if the linearized system is controllable at the origin - specifically, there exists a continuous function v such that the equation:

$$h(t) = \int_{t_0}^{t} f_x(x, u) h(\tau) d\tau + \int_{t_0}^{t} f_u(x, u) v(\tau) d\tau$$

has a solution with $h(t_1) = e$ for any $e \in \mathbb{R}^n$, then, (x, u) is a regular point for the constraint $A(x, u) = 0$.

In this case, by the Lagrange multiplier theorem, there exists a $\lambda(\cdot) \in NBV^n[t_0, t_f]$ (the space of regular functions of bounded variation) such that

$$\bar{J}(x_0, u_0, \lambda) = \int_{t_0}^{t_f} l(x_0, u_0)dt + \int_{t_0}^{t_f} d\lambda'(t) \left[x_0(t) - x_0(0) - \int_{t_0}^{t} f(x_0, u_0)d\tau \right]$$

is stationary at the optimal solution (x_0, u_0). On integration by parts, we get:

$$\bar{J}(x_0, u_0, \lambda) = \int_{t_0}^{t_f} \left[l(x_0, u_0) + \lambda'(t)(f(x, u) - \dot{x}) \right] dt + \left[\lambda'(t) \left(x_0(t) - x_0(t_0) - \int_{t_0}^{t} f(x_0, u_0) d\tau \right) \right]_{t_0}^{t_f}$$

As $x_0(t_f) = x_0(t_0) + \int_{t_0}^{t_f} f(x_0, u_0) \, dt = 0$ we are led to the conclusion that the augmented cost function

$$\bar{J}(x_0, u_0, \lambda) = \int_{t_0}^{t_f} \left[l(x_0, u_0) + \lambda'(t)(f(x, u) - \dot{x}) \right] dt$$

is stationary at the point $(x_0, u_0, \lambda) \in C^n[t_0, t_f] \times C^m[t_0, t_f] \times NBV^n[t_0, t_f]$. Therefore, in the proof of theorem 2.3, we focus attention on the augmented cost function directly.

(Proof of Theorem 2.3)

For each $\sigma \in (-\epsilon, \epsilon)$, consider the augmented cost function:

$$\bar{J}(q(\cdot, \sigma), V(\cdot, \sigma), p_1(\cdot, \sigma), p_2(\cdot, \sigma)) = \int_{t_0}^{t_f} \left(L(q, V, u) + p_1 (\dot{q} - V) + p_2 \left(\frac{DV}{dt} - f(q, V, u) \right) \right) dt, \tag{29}$$

where we have denoted $q(t, \sigma), V(t, \sigma), p_1(t, \sigma)$ and $p_2(t, \sigma)$ in the integral, by q, V, p_1, p_2 for conciseness.

Before proceeding further, we note that by the Chain Rule:

$$
\begin{aligned}
\frac{\partial}{\partial \sigma} L(q(t,0), V(t,0), u(t,0)) &= \left[L_q(q(t,\sigma), V(t,\sigma), u(t,\sigma)) \frac{\partial q}{\partial \sigma} \right. \\
&\quad + L_{V^i}(q(t,\sigma), V^i(t,\sigma) E_i(t,\sigma), u(t,\sigma)) \frac{\partial}{\partial \sigma} \left(V^i(t,\sigma) E_i(t,\sigma) \right) \\
&\quad + \left. L_{u^j}(q(t,\sigma), V(t,\sigma), u(t,\sigma)) \frac{\partial u^j}{\partial \sigma} \right]_{\sigma=0} \\
&= L_q(q, V, u)(t, 0)(\delta q(t)) + L_V(q, V, u)(t, 0) \left(\frac{DV}{d\sigma}(t) \right) \\
&\quad + L_u(q, V, u)(t, 0)(\delta u(t)).
\end{aligned} \tag{30}
$$

Similarly,

$$\frac{\partial}{\partial \sigma} f(q, V, u)(t, 0) = \frac{\partial}{\partial \sigma} f^k(q, V, u)(t, 0) E_k(t, 0)$$

14

$$= \left(f_q^k(q,V,u)(t,0)(\delta q) + f_V^k(q,V,u)(t,0)(\frac{DV}{d\sigma}(t)) \cdots \right.$$
$$\left. + f_u^k(q,V,u)(t,0)(\frac{\partial u}{\partial \sigma}(t)) \right) E_k(t,0) + f^k(q,V,u)(t,0)\frac{DE_k(t,0)}{d\sigma} \quad (31)$$

Lets consider the last term in the above equation.

$$
\begin{aligned}
f^k(q,V,u)(t,0)\frac{DE_k(t,0)}{d\sigma} &= f^k(q,V,u)(t)\nabla_{\delta q(t)}E_k(t) \\
&= f^k(q,V,u)(t)\omega_{jk}^i \delta q^j E_i \\
&= f^k(q,V,u)(t)(\omega_{kj}^i - C_{kj}^i)\delta q^j E_i. \quad (32)
\end{aligned}
$$

Now by Lebesgue's Dominated Convergence Theorem, we have:

$$
\begin{aligned}
\frac{\partial \bar{J}}{\partial \sigma}(q(\cdot,0),V(\cdot,0),p_1(\cdot,0),p_2(\cdot,0)) &= \int_{t_0}^{t_f} \frac{\partial}{\partial \sigma}\left(L(q,V,u) + p_1(\dot{q}-V) + p_2(\frac{DV}{dt}-f(q,V,u)) \right)_{\sigma=0} dt \\
&= \int_{t_0}^{t_f} \left(\frac{\partial}{\partial \sigma}L(q,V,u) + p_1(\frac{D\dot{q}}{d\sigma}) - p_1(\delta V) + p_2(\frac{D}{d\sigma}\frac{DV}{dt}) \right. \\
&\qquad \left. -p_2\frac{\partial}{\partial \sigma}f(q,V,u) \right) dt.
\end{aligned}
$$

By Lemmas 2.3, 2.4 and Equations $(30-32)$, we get:

$$
\begin{aligned}
\frac{\partial \bar{J}}{\partial \sigma}(q(\cdot,0),V(\cdot,0),p_1(\cdot,0),p_2(\cdot,0)) &= \int_{t_0}^{t_f} \big(L_q(q,V,u)\delta q + L_V(q,V,u)\delta V + L_u(q,V,u)\delta u \\
&\qquad -\frac{Dp_1}{dt}(\delta q) - p_1(\delta V) + \Sigma R(\Sigma^{-1}p_2,V)V(\delta q) - \frac{Dp_2}{dt}(\delta V) \\
&\qquad -p_2(f_q(\delta q)) - p_2(f_v\delta V) - p_2(f_u\delta u) \\
&\qquad -[\omega(f)]^*p_2(\delta q) + [C(f)]^*p_2(\delta q) \big) dt
\end{aligned}
$$

As the variations δq, δV and δu are arbitrary, subject to $\delta q(t_0) = \delta q(t_f) = \delta V(t_0) = \delta V(t_f) = 0$ we have the first two statements of the theorem.

To prove the last statement of the theorem, consider

$$
\begin{aligned}
\dot{H}(t) &= \frac{d}{dt}L(q_0,V_0,u_0)(t) - \frac{Dp_1}{dt}(V_0) - p_1(\frac{DV}{dt})(t) - \frac{Dp_2}{dt}(f(q_0,V_0,u_0))(t) \\
&\quad -p_2(\frac{D}{dt}f(q_0,V_0,u_0))(t) \\
&= L_q(q_0,V_0,u_0)(V_0) + L_V(q_0,V_0,u_0)(\frac{DV}{dt}) + L_u(q_0,V_0,u_0)(\dot{u}) \\
&\quad - \left(L_q(q_0,V_0,u_0) + \Sigma R(\Sigma^{-1}p_2,V_0)V_0 - (f_q^*(q_0,V_0,u_0) + \omega^*(f) - C^*(f))p_2 \right)(V_0)(t) \\
&\quad -p_1(f(q_0,V_0,u_0))(t) - (-p_1 + L_V(q_0,V_0,u_0) - f_V^*(q_0,V_0,u_0)p_2)(f(q_0,V_0,u_0))(t) \\
&\quad -p_2 \left(f_q(q_0,V_0,u_0)(V_0) + f_V(q_0,V_0,u_0)(\frac{DV}{dt}) + f_u(q_0,V_0,u_0)(\dot{u}) + f^i \frac{DE_i}{dt} \right) \\
&= 0,
\end{aligned}
$$

15

by Equations (15),(16); the first two statements that we just proved; and the fact that

$$< R(\Sigma^{-1}p_2, V_0)V_0, V_0 >= 0$$

by Proposition 2.2.   ∎

3 Applications

In this section, we specialize the results of the last section to obtain a formula for cubic splines on Riemannian manifolds and to rotational rigid body dynamics. We then present a numerical example where we apply a new numerical method called the *Modified Simple Shooting Method*[13] and solve the resulting two-point boundary value problem for the rotational rigid body dynamics problem.

3.1 Cubic splines on Riemannian manifolds

Here we specialize Theorem 2.3 and recover the Noakes, Heinzinger and Paden formula for cubic splines on Riemannian manifolds [5]. Let M be a parallelizable Riemannian manifold and let $q_0, q_1 \in M$, $V_0 \in T_{q_0}M$ and $V_1 \in T_{q_1}M$. Consider the problem:

$$\text{Minimize} \quad J(u(\cdot)) = \int_{t_0}^{t_f} u^2(t)dt$$

subject to:

$$\dot{q}(t) = V(t), \tag{33}$$

$$\frac{DV}{dt} = u(t) = u^i(t)E_i(t), \tag{34}$$

and

$$q(t_0) = q_0$$

$$q(t_f) = q_f$$

$$\dot{q}(t_0) = V_0$$

$$\dot{q}(t_f) = V_f \tag{35}$$

Thus we have $f(q, V, u) = u$ and Σ is the identity matrix. Therefore, there is an identification of vectors and co-vectors. Then, Theorem 2.3 asserts the existence of one-form sections $p_1(t), p_2(t)$

such that:

$$\frac{Dp_1}{dt} = R(p_2, V)V - (\omega^*(f) - C^*(f))p_2$$

$$\frac{Dp_2}{dt} = -p_1$$

$$p_2 = u,$$

where we have used the identification of vectors and co-vectors in the last equation. Thus

$$\frac{D^2V}{dt^2} = \frac{Du}{dt} = \frac{Dp_2}{dt} = -p_1$$

$$\frac{D^3V}{dt^3} = -\frac{Dp_1}{dt} = -R(p_2, V)V + (\omega^*(f) - C^*(f))p_2.$$

Let $w \in \Psi(M)$. Then

$$(\omega^*(f) - C^*(f))p_2(w) = p_{2_i}(\omega^i{}_j(p_2) - C^i{}_j(p_2))w^j$$

$$= p_{2_i}\omega^i{}_{jk}p_2^k w^j$$

$$= \omega^i{}_k(w)p_2^k p_{2_i}$$

$$= 0 \quad \text{for all} \quad w \in \Psi(M)$$

because $\omega^i{}_k = -\omega^k{}_i$ by Proposition 2.1. Therefore,

$$\frac{D^3V}{dt} + R(\frac{DV}{dt}, V)V = 0, \tag{36}$$

which is the equation for a cubic spline that was first obtained by Noakes, Heinzinger and Paden [5].

3.2 Rigid body rotation

Consider a rigid body with principal moment of inertia matrix $\mathbb{I}$ and mass m. The configuration space of a rigid body $SE(3)$ is the space of rotations given by the set of 3×3 matrices $SO(3) = \{R \mid R'R = I; \det(R) = 1. R$ is the orientation of the rigid body with respect to an earth-fixed co-ordinate system.

The angular velocity of the body in the principal axis system (called the body axis system) centered at the center of mass is defined as: $\Omega = Q'\dot{Q}$ (where $'$ denotes the matrix transpose). If we define the skew-symmetrization of $\Omega' = [\Omega_1, \Omega_2, \Omega_3]$ to be

$$\hat{\Omega} = \begin{bmatrix} 0 & -\Omega_3 & \Omega_2 \\ \Omega_3 & 0 & -\Omega_1 \\ -\Omega_2 & \Omega_1 & 0 \end{bmatrix}$$

then the Euler's equations for a rigid body can be written in a compact form as:

$$\dot{Q} = Q\hat{\Omega} \tag{37}$$

$$\dot{\Omega} - \Omega \times \mathbb{I}^{-1}\Omega = \mathbb{I}^{-1}T_e. \tag{38}$$

where T_e are the moments acting on the body expressed in the body axis system. In the following, we treat $\mathbb{I}^{-1}T_e$ as the input vector u.

We now derive the same equations using a geometrical method. Let $\{e_1, e_2, e_3\}$ form a basis for the Lie algebra $so(3)$ of the Lie Group $SO(3)$. Then one can form a parallel frame at each point $Q \in SO(3)$ given by $E_i(Q) = Qe_i$, $i = 1, 2, 3$ for the Lie group $SE(3)$. For this parallel frame $\{E_1, E_2, E_3\}$ let the structure constants C_{ij}^k for the Jacobi-Lie bracket can be obtained from those for the Lie algebra bracket via the following computation:

$$[E_i(Q), E_j(Q)]_{J-L} = [Qe_i, Qe_j]$$
$$= -Q[e_i, e_j]_{LA},$$

where $[\cdot, \cdot]_{J-L}$ denotes the Jacobi-Lie bracket of vector fields and $[\cdot, \cdot]_{LA}$ denotes the Lie algebra bracket of vectors in $so(3)$. Suppose we chose the structure constants for the Lie algebra bracket to be as in the following table.

$c_{12}^3 = \frac{1}{I_3}$	$c_{23}^1 = -\frac{1}{I_1}$	$c_{31}^2 = -\frac{1}{I_2}$	$c_{ij}^k = 0$ for all other $1 \le i, j, k \le 3$.

Then the structure constants for the Jacobi-Lie bracket of co-ordinate frame vector fields C_{jk}^i is given by:

$$C_{jk}^i = -c_{jk}^i; \quad 1 \le i, j, k \le 3.$$

The Riemannian metric is defined via the following table:

$<E_1, E_1> = \frac{1}{2}I_1$		$<E_2, E_2> = \frac{1}{2}I_2$	$<E_3, E_3> = \frac{1}{2}I_3$
$<E_i, E_j> = 0$ for all other $1 \le i, j \le 6$.			

We can compute the connection matrix for the rotating rigid body using Proposition 2.1 and Theorem 2.2. If $\{\theta^1, \theta^2, \theta^3\}$ is a set of co-vector fields on $SO(3)$ such that $\theta^i(E_j) = \delta_j^i$, then,

Equation (9) leads to:

$$d\theta^i(E_j, E_k) = E_j(\theta^i(E_k)) - E_k(\theta^i(E_j)) - \theta^i([E_j, E_k])$$
$$= -\theta^i([E_j, E_k]).$$

We now equate the left-hand-side of the above equation with that of Equation (7) and compute (by solving for the one-forms ω^k_i's) the following connection matrix:

$$\left[\omega^k_i\right] = \begin{array}{c} k \\ \downarrow \end{array} \overset{i \rightarrow}{\begin{bmatrix} 0 & \alpha\theta^3 & \beta\theta^2 \\ -\alpha\theta^3 & 0 & \gamma\theta^1 \\ -\beta\theta^2 & -\gamma\theta^1 & 0 \end{bmatrix}}, \tag{39}$$

where $\alpha = -\frac{1}{2}\left(-\frac{1}{I_1} - \frac{1}{I_2} + \frac{1}{I_3}\right)$, $\beta = -\frac{1}{2}\left(\frac{1}{I_1} - \frac{1}{I_2} + \frac{1}{I_3}\right)$ and $\gamma = -\frac{1}{2}\left(\frac{1}{I_1} - \frac{1}{I_2} - \frac{1}{I_3}\right)$. Note that the ω^k_i matrix is skew-symmetric.

Next, we compute the matrix $\nabla_E E$ using Equation (3):

$$\left[\nabla_{E_j} E_i\right] = \left[\omega^k_i(E_j)E_k\right] = \begin{array}{c} i \\ \downarrow \end{array} \overset{j \rightarrow}{\begin{bmatrix} 0 & -\beta E_3 & -\alpha E_2 \\ -\gamma E_3 & 0 & \alpha E_1 \\ \gamma E_2 & \beta E_1 & 0 \end{bmatrix}}, \tag{40}$$

Using the above matrix, we can obtain an expression for the co-variant derivative of a vector field on $SO(3)$ with respect to another vector field on $SO(3)$. Let $\Omega, \xi \in \Psi(SO(3))$ and $q(t)$ be the curve obtained by solving the system $\dot{q}(t) = \Omega$; $q(0) = q_0$; $t \in [0, 1]$. Then along the curve $q(\cdot)$, we compute:

$$\nabla_\Omega \xi = \begin{bmatrix} \dot{\xi}^1 \\ \dot{\xi}^2 \\ \dot{\xi}^3 \end{bmatrix} + \begin{bmatrix} \alpha\xi^2\Omega^3 + \beta\xi^3\Omega^2 \\ -\alpha\xi^1\Omega^3 + \gamma\xi^3\Omega^1 \\ -\beta\xi^1\Omega^2 - \gamma\xi^2\Omega^1 \end{bmatrix}$$
$$= \dot{\xi} + \xi \times \Theta\Omega, \tag{41}$$

where

$$\Theta = \begin{bmatrix} \gamma & 0 & 0 \\ 0 & -\beta & 0 \\ 0 & 0 & \alpha \end{bmatrix}. \tag{42}$$

19

In particular,

$$\frac{D\Omega}{dt} = \nabla_\Omega \Omega = \dot{\Omega} + \Omega \times \Theta\Omega. \tag{43}$$

It is interesting to note that even though $\Theta \neq \mathbb{I}^{-1}$, we still have $\Omega \times \Theta\Omega = \Omega \times \mathbb{I}^{-1}\Omega$!

We compute the curvature tensor $R(\xi, \Omega)\Omega$ using Proposition 2.3. The curvature two-form Ω^k_i matrix turns out to be (we apologize for the the use of Ω to denote both angular velocities and the curvature two-form, but the curvature two-form only appears on this page):

$$\left[\Omega^k_i \right] = \begin{matrix} \\ k \\ \\ \downarrow \end{matrix} \overset{i \rightarrow}{\begin{bmatrix} 0 & (\frac{\alpha}{I_3} + \beta\gamma)\theta^1 \wedge \theta^2 & (\frac{\beta}{I_2} + \alpha\gamma)\theta^3 \wedge \theta^1 \\ -(\frac{\alpha}{I_3} + \beta\gamma)\theta^1 \wedge \theta^2 & 0 & (\frac{\gamma}{I_1} + \alpha\beta)\theta^2 \wedge \theta^3 \\ -(\frac{\beta}{I_2} + \alpha\gamma)\theta^3 \wedge \theta^1 & -(\frac{\gamma}{I_1} + \alpha\beta)\theta^2 \wedge \theta^3 & 0 \end{bmatrix}} \tag{44}$$

As

$$R(\xi, \Omega)E_j = \Omega^i{}_j \, (\xi, \Omega)E_i,$$

we have:

$$R(\xi, \Omega)\Omega = \Omega \times \Lambda_1(\xi \times \Omega), \tag{45}$$

where

$$\Lambda_1 = \begin{bmatrix} \frac{\gamma}{I_1} + \alpha\beta & 0 & 0 \\ 0 & -\left(\frac{\beta}{I_2} + \alpha\gamma\right) & 0 \\ 0 & 0 & \frac{\alpha}{I_3} + \beta\gamma \end{bmatrix}.$$

Finally, we compute the $(\omega^*(f) - C^*(f))p_2$ term that appears in Theorem 2.3:

$$(\omega^*(f) - C^*(f))p_2 = p_{2_i}(\omega^i{}_j \, (u^k E_k)\sigma^j - C^i{}_j \, (u^k E_k)\sigma^j); \quad 1 \leq i, j, k \leq 3.$$

After extensive computations, the right hand side turns out to be:

$$(\omega^*(f) - C^*(f))p_2 = \Theta \begin{bmatrix} u_1 \\ u_2 \\ u_3 \end{bmatrix} \times \begin{bmatrix} p_{2_1} \\ p_{2_2} \\ p_{2_3} \end{bmatrix}. \tag{46}$$

Theorem 2.3 postulates the existence of one-forms $p_1(t)$ and $p_2(t)$ for $t \in [t_0, t_f]$. Define the vector fields $\xi(\cdot)$ and $\eta(\cdot)$ along $q(\cdot)$, via the identification $\xi(t) = \Sigma^{-1}p_1(t)$ and $\eta(t) = \Sigma^{-1}p_2(t)$. Then, we can write the necessary conditions in terms of (Q, Ω, ξ, η). The full set of equations are then:

$$\dot{Q} = Q\hat{\Omega} \tag{47}$$

$$\dot{\Omega} = -\Omega \times \Theta\Omega + \eta \tag{48}$$

20

$$\dot{\xi} = -\xi \times \Theta\Omega + \Omega \times \Lambda_1(\eta \times \Omega) - \Theta\Sigma^{-1}\eta \times \eta \tag{49}$$

$$\dot{\eta} = -\eta \times \Theta\Omega - \xi, \tag{50}$$

with $(Q, \Omega)(t_0)$ and $(Q, \Omega)(t_f)$ specified.

4 Numerical Experiments

The equations (47–50) constitute a two-point boundary value problem. We used a modified shooting method technique [13] to numerically solve for the unknown "Lagrange multipliers" (ξ, η) at initial time.

The first equation (47) is a matrix equation that we integrated using the well-known Rodriguez's formula [14]:

$$Q(t+h) = Q(t) \left(I + \frac{\hat{\Omega}}{\|\Omega\|} \sin(\|\Omega\|h) + \frac{\hat{\Omega}^2}{\|\Omega\|^2}(1 - \cos(\|\Omega\|h)) \right), \tag{51}$$

where h is the time-step for integration. This results in a Q matrix that "stays" on the group $SO(3)$ at each time-step.

The moments of inertia constants for the numerical simulation were chosen arbitrarily to be: $I_1 = 10$; $I_2 = 5$; $I_3 = 2.5$. The initial time t_0 was set to 0 and the final time t_f was chosen to be 1. The initial states $(Q, \Omega)(0)$ and the desired final states (Q_{des}, Ω_{des}) at $t = 1$, were chosen using the pseudo-random number generating program in MATLAB. Their values for a simulation run are listed in the following table (to 3 significant digits).

$Q(0)$			$\Omega(0)$
0.572	0.817	0.079	0.950
−0.783	0.572	−0.246	4.337
−0.246	0.079	0.966	7.0923

$Q_{des}(1)$			$\Omega_{des}(1)$
−0.268	0.963	−0.011	1.901
−0.838	−0.227	0.497	8.673
0.476	0.142	0.868	4.185

The initial value for the co-states $(\xi, \eta)(0)$ was also chosen using the pseudo-random number generating program in MATLAB and it turned out to be

$\xi(0)$	$\eta(0)$
0.116	0.034
0.078	0.192
0.369	0.471

The modified simple shooting method involves the choice of a continuous, time-parametrized reference path that connects the initial and final points. For $t \in [0, 1]$, we picked the reference path to be:

$$Q_{ref}(t) = Q(0)\exp(\hat{\phi}t), \quad \text{where} \quad \hat{\phi} = \ln(Q(0)^{-1}Q_{des})$$
$$\Omega_{ref}(t) = \Omega(0) + (\Omega_{des} - \Omega(0))t.$$

The equations were integrated in the forward direction until at some time $t \in (0, 1]$, we had

$$\|q(t) - q_{ref}(t)\| = 100 * \|Q(t) - Q_{ref}(t)\|_M + \|\Omega(t) - \Omega_{ref}(t)\| \geq 200. \tag{52}$$

At this point, the initial guesses for the Lagrange multipliers $(\xi, \eta)(0)$ were updated via the modified Newton's method until $\|q(t) - q_{ref}(t)\| \leq 0.5$. Then the equations are integrated forward again until the inequality in (52) is satisfied. The vector $(\xi, \eta)(0)$ is updated as before and the iteration is repeated. The orientation matrix norm was calculated as: $\|Q\|_M = \|\ln Q\|$ where $\|\cdot\|$ is the standard matrix norm. The norm on the orientation matrix was multiplied by 100 so that the final orientation is met accurately. The time step for the integration was 0.02 seconds. The CPU time taken for the computations in a MATLAB environment, running on a 1.8 GHz PC was 32.94 seconds.

The solution of the two-point boundary value problem led to the following final states at $t = 1$:

$Q(1)$			$\Omega(1)$
-0.268	0.963	-0.011	1.901
-0.838	-0.227	0.497	8.673
0.476	0.142	0.868	4.185

The program converged to the following value of the co-states at $t = 0$.

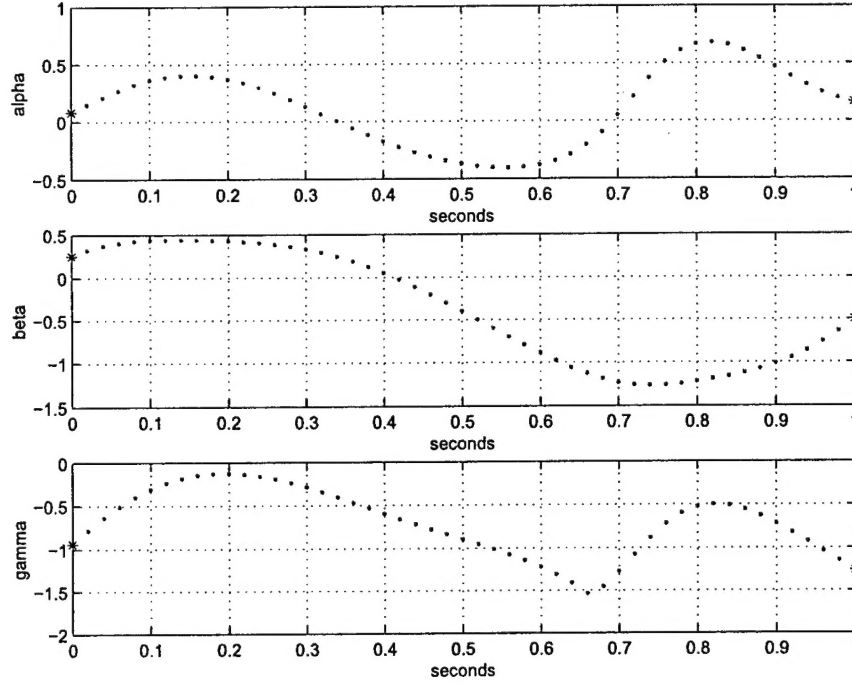| $\xi(0)$	$\eta(0)$
-56.132	17.794
-84.811	47.803
-65.834	33.502

Figure 2: Plot of ZYX Euler angles.

Figures 2 − 4 show the results of the simulation. The ZYX Euler angles in Figure 2 was computed according to:

$$\begin{aligned}
\beta &= -\sin^{-1}(Q_{31}) \\
\alpha &= \sin^{-1}\left(\frac{Q_{32}}{\cos(\beta)}\right) \\
\gamma &= \sin^{-1}\left(\frac{Q_{21}}{\cos(\beta)}\right).
\end{aligned}$$

Though $\gamma(t)$ in the figure seems to be not continuously differentiable at a point, one can clearly see from the plots of $Q_{31}(t), Q_{32}(t)$ and $Q_{21}(t)$ in Figure 3 that they are indeed continuously differentiable.

5   Conclusion

In this paper, we have made three new contributions. Firstly, we have derived first order necessary conditions for an optimal control problem on a parallelizable Riemannian manifold. These equations specialize to those of cubic splines on Riemannian manifolds that were first discovered by
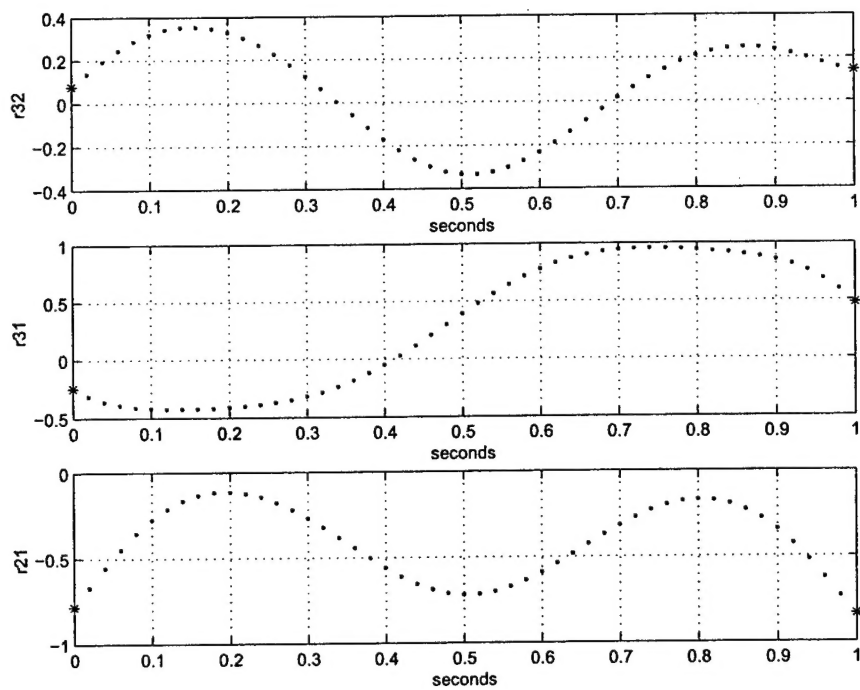
23

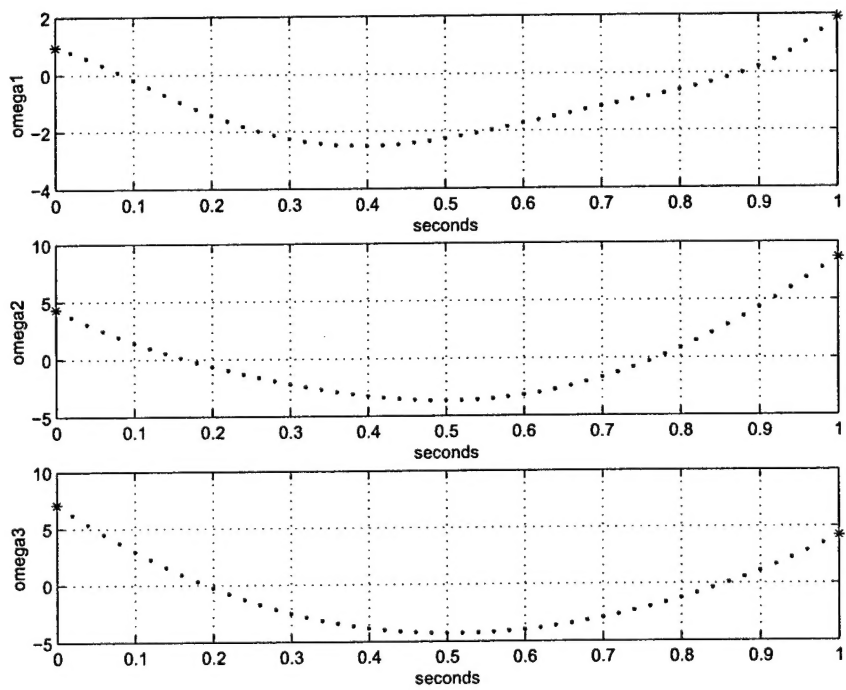Figure 3: Plot of orientation matrix entries.

Figure 4: Plot of angular velocities.

24

Noakes, Heinzinger and Paden. Secondly, we have specialized the equations to a rigid body rotation problem. Thirdly, we have presented the results of numerical experiments where we solve the two point boundary value problem resulting from the necessary conditions. We used a modified simple shooting method for the numerical solution of the problem and the results indicate the feasibility of online implementation of the path planning problems.

References

[1] P. Crouch and F. S. Liete, "The dynamic interpolation problem: On Riemannian manifolds, Lie groups, and symmetric spaces," *Journal of Dynamical and Control Systems*, vol. 1, no. 2, pp. 177–202, 1995.

[2] P. Crouch, F. S. Liete, and M. Camarinha, "A second order Riemannian varational problem from a Hamiltonian perspective." Private Communication, 2001.

[3] P. Crouch, M. Camarinha, and F. S. Leite, "Hamiltonian approach for a second order variational problem on a Riemannian manifold," in *Proceedings of CONTROLO'98, 3rd Portuguese Conference on Automatic Control*, pp. 321–326, Sept. 1998.

[4] H. Sussmann, "An introduction to the coordinate-free maximum principle," in *Geometry of Feedback and Optimal Control* (B. Jakubczyk and W. Respondek, eds.), pp. 463–557, Marcel Dekker, New York, 1997.

[5] L. Noakes, G. Heinzinger, and B. Paden, "Cubic splines on curved spaces," *IMA Journal of Mathematical Control and Information*, vol. 6, pp. 465–473, 1989.

[6] T. Frankel, *The Geometry of Physics: An Introduction.* Cambridge University Press, 1998.

[7] P. Petersen, *Riemannian Geometry.* Springer Verlag, 1998.

[8] W. M. Boothby, *An introduction to Differential Geometry and Riemannian Manifolds.* Academic Press, 1975.

[9] M. P. D. Carmo, *Riemannian Geometry.* Birkhauser, 1993.

[10] W. Klingenberg, *Riemannian Geometry.* Walter de Gruyter, Berlin, 1982.

[11] S. Helgason, *Differential Geometry, Lie Groups and Symmetric Spaces.* Academic Press, 1978.

[12] D. G. Luenberger, *Optimization by Vector Space Methods.* John Wiley and Sons, 1969.

[13] R. Holsapple, R. Venkataraman, and D. Doman, "A modified simple shooting method for solving two point boundary value problems." Will appear in the proceedings of the IEEE Aerospace Conference, Big Sky, MT, March 2003.

[14] R. M. Murray, Z. Li, and S. S. Sastry, *A Mathematical Introduction to Robotic Manipulation.* CRC Press, 1994.